DOCUMENT RESUME

ED 410 285 TM 027 108

AUTHOR Reckase, Mark D.

TITLE Constructs Assessed by Portfolios: How Do They Differ from Those Assessed by Other Educational Tests.

PUB DATE Mar 97

NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).

PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Cluster Analysis; College Entrance Examinations; Content Analysis; *Educational Assessment; English; Item Response Theory; Language Arts; Mathematics; Multiple Choice Tests; *Portfolio Assessment; Portfolios (Background Materials); Sciences; *Standardized Tests; Student Evaluation; *Test Results; Writing (Composition)

IDENTIFIERS ACT Assessment; *Constructs

ABSTRACT

Portfolio assessment has become a widely cited technique for obtaining a rich demonstration of the skills that students have acquired, and it is often implied that portfolio assessment provides different types of information from that obtained from standardized tests. This paper provides support for the contention that portfolio assessments do indeed provide information about student capabilities that is different from that available from standardized tests. This support is given in the context of a portfolio assessment model called PASSPORT designed for use with students in grades 9 through 12. PASSPORT is a portfolio assessment system for collecting and presenting evidence of student knowledge and skills in the areas of English language arts, mathematics, and science. Students produce portfolios of their work through a loose set of guidelines. Each portfolio contains five work samples and a cover letter that explains why samples were included and what they tell about the student's capabilities. To compare the content assessed by PASSPORT and the American College Testing Program (ACT) Assessment, a multidimensional item response theory model (MIRT) was applied to the mathematics and language arts scores from the portfolio and the multiple choice test. The MIRT analysis was followed by a cluster analysis of similarities of the dimensions assessed by the items. Comparison of the content and cluster analyses shows clear differences in the types of skills and knowledge that are assessed. The portfolio provides evaluations of student performance on major writing tasks and on mathematics skills related to data analysis and problem solving, while the standardized test provides information about the details of the writing process and the rules for manipulating mathematical expression. (Contains two figures, six tables, and nine references.) (SLD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Mark Reckase
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

# Constructs Assessed by Portfolios: How Do they Differ from those Assessed by other Educational Tests[1]

Mark D. Reckase
ACT, Inc.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED 410 285

Portfolio assessment has become a widely cited technique for obtaining a rich demonstration of the skills that students have acquired through their interaction with the educational system (e.g., Bridgeman, Chittenden & Cline, 1995; Brookhart & Masciola, 1996; Glazer & Brown, 1993). It is also often implied that portfolio assessment provides different types of information than that obtained from standardized tests. For example, in a discussion of portfolio assessment, Glazer & Brown (1993) indicate:

> "Using formal and informal tests that do not reflect students' authentic reading and writing may solve problems related to *reliability* across students. This practice, however, creates even more serious *validity* problems. Looking at products while ignoring the processes students use to create them provides distorted and incomplete pictures of students' abilities" (pp. 15-16).

They imply that portfolios and other "authentic" measures will provide the information on process that is lacking in other procedures. But is there any evidence that this is really the case? What do portfolio assessments measure?

TM027108

[1]Paper presented at the annual meeting of the American Educational Research Association, Chicago, March, 1997.



BEST COPY AVAILABLE

In a similar vein, Myers and Pearson (1996) suggest that standardized tests are "narrowly conceived" and that they measure "general knowledge" rather than more appropriate skills. While they do acknowledge the value of multiple-choice tests, the tests are considered "the enemy with whom we were doing battle." The use of multiple-choice tests is relegated to getting information about performance of "students at the low end of the scale." The implication is that multiple-choice tests measure different skills than portfolios and that these skills are at the lower end of the skills taught in the curriculum.

Callahan (1995) expresses similar beliefs for the area of writing. Multiple-choice tests "are said to lack validity primarily because they do not involve the production of writing samples." Portfolios "give a broader picture of what a writer can do, over a longer period of time, under varied circumstances, in response to a number of writing opportunities."

Many authors suggest that portfolios assess different skills than other types of tests, a review of the literature found very few studies that documented such differences. A study by Brookhart and Masciola (1996) reported an analysis of 30 sets of 8th grade writing portfolio scores from a school district in Pennsylvania. Their goal was to determine whether the portfolios provided information that was different than that provided by standardized tests. Using factor analysis, they found that scores from the portfolios tapped both a classroom work/achievement factor and a verbal ability factor while the standardized tests were related only to the verbal ability factor.

A second study by Bridgeman, Chittenden, and Cline (1995) found that results from a literacy portfolio were more highly related to performance on a standardized test administered later than were scores from a concurrently administered standardized test. These results were for 251 first grade students from one school district. They indicate that the portfolios "provided a richer source of diagnostic information" than the standardized multiple-choice tests.

One notable point about these studies is that the definition of a portfolio was usually made explicit. In the Brookhart and Masciola (1996) study, the table of contents for the portfolio was listed, but the conceptual framework for the portfolio was not described. It was never stated whether the portfolio was designed to showcase students' best work, or to provide a sampling of typical work. It was implied, but not explicitly stated, that the portfolio represented student work relative to school curriculum objectives and state learning outcomes, but the relationship of the final portfolio design to those educational goals was never made clear.

Similarly, the description of the portfolio in Bridgeman, Chittenden, and Cline (1995) included a table of contents and provided the statement that the portfolio "carries direct evidence of the child's initial strategies and attempts to 'make sense of print.'" However, no conceptual definition of the portfolio was given.

The conceptual basis for the portfolio being developed by the New Standards Project (Myers & Pearson, 1996) is even less clear. They indicate that the portfolio work reflects "community-specific knowledge," (p. 13) but never give a hint as to what that is. They indicate that the portfolio has a tentative outline of content, but no detail is provided.

Thus, there is some weak support for the contention that portfolios provide different kinds of information than standardized tests, but the definition of portfolio assessment is not clear, and more research on the subject is clearly needed.

The purpose of this paper is to provide some additional support for the contention that portfolio assessments provide information about student capabilities that is different than that available from standardized tests. This support is given in the context of a portfolio assessment model called PASSPORT (ACT, 1996) that is designed for use with students in grades nine through twelve. A formal definition of the type of portfolio being used is provided along with an analysis of the skills assessed by PASSPORT as compared to a standardized achievement test, the ACT Assessment (ACT, 1988). More specifically, the content specifications for the two assessment programs and the dimensional structure of assessments for the same curriculum areas will be compared.

## The PASSPORT System

PASSPORT is a portfolio assessment system for use in collecting and presenting evidence of students' knowledge and skills in the areas of English language arts, mathematics, and science. The system is designed for use with students from grades nine through twelve. Using the system, students produce portfolios of their work according to a loose set of guidelines. Each portfolio contains five examples of student work, called work samples, and a cover letter that explains to the reader of the portfolio the reasons the work samples were selected for inclusion and what the samples tell about the student's capabilities.

The conceptual framework for the PASSPORT portfolios is the definition developed by Meyer, Schuman, and Angello (1990):

> A portfolio is a purposeful collection of student work that tells the story of the student's efforts, progress, or achievement in given areas. This collection must include:
>
> (1) student participation in selection of portfolio content;
>
> (2) the guidelines for selection;
>
> (3) the criteria for judging merit; and
>
> (4) evidence for student self-reflection.

As specified by the definition, PASSPORT supports portfolio use by providing the guidelines for selection, the criteria for judging merit, and a system that leads the student through the process of selection of materials and reflecting on the qualities of the materials.

The feature of the PASSPORT System that defines the structure and content of the portfolios is a menu of Work Sample Descriptions. Work Sample Descriptions are general explanations of the requirements for an entry into the portfolio. Work Sample Descriptions are not prompts, but rather are lists of desired characteristics for a portfolio entry. Tables 1 and 2 provide lists of the Work Sample Descriptions for the English language arts and mathematics portfolios, respectively.

---

Insert Tables 1 and 2 about here

---

To define a PASSPORT portfolio, five Work Sample Descriptions are selected from the menu and students select work from their regular class assignments that they believe fit the requirements of the Work Sample Descriptions. Those selections, plus a self-reflective cover letter make up the portfolio.

Scoring rubrics have been developed for submissions related to each Work Sample Description and for the portfolio as a whole. Each work sample receives a score on a 1 to 6 scale, and the portfolio as a whole receives a rating related to one of four categories.

The menu system for PASSPORT allows the portfolios to be customized to match a school's instructional programs and the interests of the students. The use of the menu system also allows each student to produce a somewhat different portfolio, although most students' portfolios have at least one Work Sample Description in common with that of another student. Thus, the "test" defined by the PASSPORT portfolio may vary from student to student.

## Content Analysis

The content assessed through the use of PASSPORT portfolios is defined by the selection of Work Sample Descriptions from the menu for the curriculum area of interest. As indicated above, this results in a portfolio, and in assessment content, that is somewhat different for different students. In contrast, the test specifications for a standardized test provides more structure and comparability across forms than does the portfolio. However, there is still quite a bit of variation in the selection of test items for equivalent forms of a test.

To make the comparison of content very concrete, the content of a typical portfolio will be compared to the content of a standardized test in the same curriculum area. For this paper, the ACT Assessment Mathematics and English Tests will be used for comparison purposes.

**English**

A typical PASSPORT portfolio in the English language arts area consists of work related to the following five Work Sample Descriptions: Analysis/Evaluation, Imaginative Writing, Response to a Literary Text, Writing about Out-of-Class Reading, and Writing about Values/Issues/Beliefs. A portfolio structured around these work samples will allow a student to demonstrate different types of writing and logical organizations for the writing. The Imaginative Writing piece provides an opportunity to show creativity and a personal voice. While it is not evident from the short titles for the Work Sample Descriptions, the scoring rubrics include consideration of mechanics, word choice and sentence structure to the extent that they promote understanding by the reader and do not interfere with comprehension. The emphasis in scoring, however, is on the clarity and effectiveness of presentation to an appropriate audience.

Note that this portfolio structure did not include Business and Technical Writing or Persuasive Writing so it did not cover the breadth of the domain covered by the menu of

Work Sample Descriptions. Instead, it emphasizes skills related to reacting to outside information.

In contrast to the portfolio content structure, the ACT Assessment English Test has a content structure that is summarized in Table 3. The test is a simulation of the editing task that is often required to improve second or third drafts of an essay. The successful examinee must detect errors in usage and mechanics and select appropriate fixes. They must also determine if material is redundant or superfluous, is appropriately organized, and is consistent in style. If problems are detected, improvements must be selected.

---

Insert Table 3 about here

---

A comparison of these two sets of specifications identifies some clear distinctions. First, the portfolio requires production of work rather than editing of an existing work. Production of work reveals the student's style and organization, the command of some mechanics, and the selection of approach to a topic. But, production also allows a student to avoid punctuation, vocabulary, sentence structures, etc. that have not been mastered.

Second, the ACT Assessment English Test is roughly equally balanced between usage/mechanics and rhetorical skills. The portfolio scoring places much less weight on

usage/mechanics. The multiple-choice test requires students to attend to many different types of writing problems, some of which might be avoided if the students were writing their own essays.

**Mathematics**

A typical PASSPORT portfolio in mathematics is structured around the following five Work Sample Descriptions: Analyzing Data, Challenging Problems, From Your Own Experience, Multiple Methods, and Technology. Note that none of these Work Sample Descriptions require a specific type of content such as geometry or algebra. They are generic in their descriptions and apply equally well to the content of any mathematics class. The rubrics for scoring the work samples bring in the concepts of correctness and completeness of work, the difficulty of the problem that is attempted, and the clarity of communication to the reader of the portfolio.

In contrast to the portfolio content description, the ACT Assessment Mathematics Test uses areas of study of mathematics as its structural framework (see Table 4). While problem solving and analysis are required of many of the items, reporting of the results is by subject matter area rather than by the cognitive approach to the items.

As with the English language arts portfolio, the mathematics portfolio allows students to avoid areas of weakness through the selection of work samples. Knowledge of

trigonometry facts will not be assessed if the student chooses to exclude work requiring trigonometry from the portfolio. Thus, there is a clear difference in breadth of content coverage. On the other hand, the portfolio emphasizes solution strategies and presentation of mathematical information, areas that it is not possible to assess using the multiple choice format. The multiple-choice test requires the student to confront all the areas in the table of specifications.

Comparison of the content specifications for the two types of instruments is informative, but are the differences that are suggested by the logical comparisons confirmed by statistical analyses? That is the topic of the next section of the paper.

## Multidimensional Structure

In order to compare the content assessed by PASSPORT and the ACT Assessment, both instruments were analyzed by procedures that provided information about the fine grained dimensional structure of the tests. The procedures are described in some detail in Miller (1997) so they will not be repeated here. The analyses consist of the application of a multidimensional IRT model to the scores of the assigned work samples for the portfolios or the item response data for the multiple-choice tests. The MIRT results were followed by a cluster analysis of the similarities of the dimensions assessed by the items. The results of these cluster analyses are tree diagrams that show which items or work samples are measuring

approximately the same set of dimensions. The tree diagrams for the PASSPORT Language Arts and Mathematics portfolios are given in Figures 1 and 2. The tree diagrams for the ACT Assessment English and Mathematics Tests are too extensive to be reproduced here, but the results are summarized in Tables 5 and 6 using the names given to the clusters.

---

Insert Figures 1 and 2 about here

---

For assessment tasks to cluster together in these analyses, there must be variation in student performance on the task and the level of performance on different tasks must vary for different students. If there is no variation in the sample of students on a particular skill, for example if all students have mastered the skill, it will not show up as a dimension or as a cluster.

---

Insert Tables 5 and 6 about here

---

**English**

The PASSPORT English Language Arts clustering shows a number of clear groupings of Work Sample Descriptions. At the top of the diagram are Writing about Out-of-Class Reading and Research/Investigative Writing, both of which require students to write about information they have acquired outside the classroom. Business and Technical Writing tends to include business letters as work samples. These letters are usually about events that occurred outside of class, such as a letter complaining about a defective product.

The second cluster, Proposing a Solution and Explanatory Writing relates to tasks that tend to have an organized step by step structure. The focus of the submissions is on the logical clarity of the ideas that are presented.

The third and largest cluster has to do with analysis of issues and persuading others that the analysis is correct. Finally, the fourth cluster, including Imaginative Writing and Relating a Personal Experience, provides work samples that are typically narrative and informal, and that allow students to consider emotion and personal voice.

Note that the clustering tends to fall along major types of writing such as expository, persuasive, and narrative, and according to the type of organizational structure that is typically used for each type of assignment.

In contrast to the PASSPORT clusters, the ACT Assessment English clusters focus on clarifying the meaning of text, correcting punctuation problems, determining what to add or delete from the essay, and selecting appropriate language to match the situation. this is not surprising since the test is designed to determine whether students can edit an essay to improve its quality.

**Mathematics**

PASSPORT Mathematics shows three fairly clear clusters of Work Sample Descriptions. The first cluster including Analyzing Data, Consumer Beware, and From Your Own Experience. The work samples that students submit for these areas typically include the analysis of data from outside of school activities. For example, a "From Your Own Experience" activity submitted by one student was an analysis of the cost of music from different music clubs.

The second cluster of Work Sample Descriptions focuses on the use of a number of different methods to solve problems. The methods are compared or analyzed to determine how they are related. The third cluster consists of Work Sample Descriptions that require challenging problems. Technology is usually applied to challenging problems that can not be solved in other ways. The problems selected from another class also tend to be very challenging.

The ACT Assessment Mathematics Clusters focus more on problem types and specific skills or knowledge than the portfolio based clusters. Several clusters relate to the rules needed to manipulate mathematical objects and specific mathematical forms. The clusters also tend to be organized according to subject matter areas such as geometry and trigonometry rather than according to the major areas listed for the portfolios.

## Summary and Conclusion

The comparison of the content and cluster analysis of the portfolio assessment and the standardized achievement test show clear differences in the types of skills and knowledge that are assessed. The portfolio assessment provides evaluations of student performance on major types of writing tasks, while the multiple choice test provides information about the details of the writing process. Similarly, the mathematics portfolio provides information about skills related to data analysis, working on challenging problems, and comparing methodologies, while the standardized test has more of a subject matter focus with particular emphasis on rules for manipulating mathematical expressions.

These analyses and the initial content analysis show that there are fairly clear differences in the skills and knowledge assessed by these two types of assessment tools. This is not to say that one is better than the other, only that they are different. It would seem that the particular method to be used should depend on the purposes of the assessment and the

types of information needed. The multiple-choice test provides more control over the skills that are assessed, requiring students to address all skill areas even if they do not feel confident about their knowledge about them. The portfolio model lets students and teachers customize the assessment to the local curriculum. Students can work around skill areas where their skills are lacking. Depending on the particular information needs, one or the other of the two methods can be selected, or a complete assessment of skills can be obtained by using both methods.

The cluster analysis of the portfolio Work Sample Descriptions shows that the skills being assessed are fairly global, including general writing types and major mathematics topics. Such assessments do not give detailed diagnostic information about specific writing or mathematics techniques. The standardized multiple choice test seems more appropriate for fine grained analysis of skills.

Overall, the two assessment methods seem complimentary to each other. On a content basis, neither is more desirable than the other. Each provides valuable information about the skills that students have acquired.

## References


ACT (1988). *ACT assessment program technical manual.* Iowa City, IA: ACT, Inc.

ACT (1996). *PASSPORT: ACT's portfolio system—facilitator's manual.* Iowa City, IA: ACT, Inc.

Brookhart, S. M. & Masciola, D. A. (1996, April). *The validity and reliability of portfolio assessment of eighth grade language arts students.* Paper presented at the annual meeting of the National Council of Measurement in Education, New York.

Bridgeman, B., Chittenden, E. & Cline, F. (1995). *Characteristics of a portfolio scale for rating early literacy* (MS# 94-08). Princeton, NJ: Educational Testing Service.

Callahan, S. (1995). Portfolio expectations: Possibilities and limits. *Assessing Writing, 2*(2), 117-151.

Glazer, S. M. & Brown, C. S. (1993). *Portfolios and beyond: Collaborative assessment in reading and writing.* Norwood, MA: Christopher-Gordon.

Meyer, C., Schuman, S., & Angello, N. (1990, September). *NWEA white paper on aggregating portfolio data.* Lake Oswego, OR: Northwest Evaluation Association

Miller, T. R. (1997, March). *Scaling issues for large-scale portfolio assessment.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Myers, M. & Pearson, P. D. (1996). Performance assessment and the literacy unit of the New Standards Project. *Assessing Writing, 3*(1), 5-29.

**Table 1**

**English Language Arts Work Sample Descriptions**

| English Language Arts Work Sample Descriptions | |
|---|---|
| ***Analysis/Evaluation*** | Analyze or evaluate different aspects or parts of a subject, object, or idea. |
| ***Business & Technical Writing*** | Demonstrate the ability to compose business or technical writing. |
| ***Evaluation of Print or Electronic Media*** | Understand and evaluate print or electronic media through writing. |
| ***Explanatory Writing*** | Explain a process or concept to another person through writing. |
| ***Imaginative Writing*** | Create a short story, poem, or play. |
| ***Persuasive Writing*** | Demonstrate the ability to persuade another person to change an opinion, belief, or behavior. |
| ***Proposing a Solution*** | Define a problem and offer a plausible solution. |
| ***Relating a Personal Experience*** | Relate a person experience which might involve a significant event, an important relationship, or a memorable place. |
| ***Research/Investigative Writing*** | Research a subject, gather and organize material, and present it clearly with well-documented sources. |
| ***Response to a Literary Text*** | Respond thoughtfully and insightfully to a literary text. |
| ***Writing About Out-of-Class Reading*** | Share a personal response to a text read outside of class. |
| ***Writing About Uses of Language*** | Explain the various uses and purposes of language. |
| ***Writing About Values/Issues/Beliefs*** | Describe or explain clearly a current cultural, educational, religious, social, or intellectual value, issue or belief that you hold and analyze how it affects you and/or others in society. |
| ***Writing a Review of Visual or Performing Arts*** | Offer a personal response to some aspect of the visual or performing arts, along with an analysis and recommendation. |

**Table 2**

**Mathematics Work Sample Descriptions**

| Mathematics Work Sample Descriptions | |
|---|---|
| ***Analyzing Data*** | Analyze and interpret data that has been collected by someone else and has been presented to you in the form of a chart, table, or graph. |
| ***Another Class*** | Use of math in some other class besides your math class. |
| ***Challenging Word Problem*** | Select a *challenging* word problem. |
| ***Collecting and Analyzing Data*** | Ability to *collect your own data*, to analyze and interpret it, and to present both the data and the conclusions you draw from it in a way that is understandable to an appropriate audience. |
| ***Comparing Notions*** | Ability to compare two mathematical notions. |
| ***Connections*** | Ability to connect two or more branches of mathematics. |
| ***Consumer Beware*** | Decide what decision the consumer should make by doing the math necessary for the consumer to be able to make the decision. |
| ***From Your Own Experience*** | Select a real-world problem *from your own experience.* |
| ***Logical Argument*** | Ability to argue logically in a mathematical context. |
| ***Multiple Methods*** | Ability to approach a problem in many different ways. |
| ***Technology*** | Tell what technological tool you used, why you chose it, how you used it, and in which step; tell how the tool made the solution of problem manageable by explaining the difficulties you would have had trying to solve the problem without it. |

**Table 3**

**Specifications for the ACT English Test**

| Content/Skills | Number of Items |
|---|---|
| Usage/Mechanics | 40 |
| Punctuation | (10) |
| Grammar & Usage | (12) |
| Sentence Structure | (18) |
| Rhetorical Skills | 35 |
| Strategy | (12) |
| Organization | (11) |
| Style | (12) |
| Total | 75 |

**Table 4**

**ACT Assessment Mathematics**
**Domain Specifications and Item Distribution**

| Content Areas | Cognitive Level | | | Number of Items |
|---|---|---|---|---|
| | Basic Skills | Applications | Analysis | |
| Pre Algebra | | | | 14 |
| Elementary Algebra | | | | 10 |
| Plane Geometry | | | | 14 |
| Coordinate Geometry | | | | 9 |
| Intermediate Algebra | | | | 9 |
| Trigonometry | | | | 4 |
| Number of Items | 20 | 32 | 8 | 60 |

**Table 5**

**AAP English Clusters**

| |
|---|
| Simplify Phrases<br>Clarify Awkward Phrasing<br>Rearrange Word Order to Clarify Meaning |
| Delete Punctuation that Disrupts Sentence Flow<br>Punctuation to Avoid Run-on Sentences<br>Set Word or Phrase off with Commas<br>Tricky Punctuation |
| Judge Relevance to Passage<br>Text Correct as is<br>Appropriate Addition to Passage |
| Word Choice--Commonly Misused Words<br>Select Proper Verb Form<br>Word, Phrase, or Sentence to Match Logical of Passage |

**Table 6**

**AAP Mathematics Clusters**

| |
|---|
| Word Problems--% and Fractions<br>Arithmetic Word Problems<br>Geometry Word Problems |
| Rules for Manipulating Symbols<br>Signed Number Arithmetic<br>Coordinate Systems |
| Applications of the Pythagorean Theorem<br>Absolute Value<br>Equation of a Line |
| Abstract Math Puzzles<br>Advanced Problems |
| Trigonometry |